ECON 5350 Final Exam – Fall 2014

Consider the following age-earnings model:

$$Wage_i = \beta_1 + \beta_2 Age_i + \epsilon_i \tag{1}$$

where $\epsilon_i \sim i.i.d.\ N(0, \sigma^2)$ and $i = 1, ..., N$. Wages are measured in dollars per hour, and age is measured in years.

1. (25 pts) **Estimation I**. Derive the ordinary least squares (OLS) estimator $b = (b_1, b_2)'$ in either matrix or summation form. Prove that it is an unbiased and consistent estimator of β, highlighting the necessary assumptions.

2. (25 pts) **Estimation II**. Assume you are given the sample of data $Wage = (10, 5, 18, 12, 15)$ and $Age = (40, 20, 60, 35, 45)$. Calculate the OLS estimates and the coefficient of determination, R^2. Calculate the residuals and plot them against age. Is there any evidence of heteroscedasiticity? Assuming there is, propose an efficient estimation procedure.

3. (25 pts) **Hypothesis Testing**. Carefully describe three equivalent ways to test the hypothesis $\beta_2 = 0.5$ (there is no need to actually calculate the numerical value of the test statistic, just carefully describe the procedure). The estimate of β_2 is positive. Does this imply that the p-value for the null hypotheses $\beta_2 = 0.5$ and $\beta_2 \leq 0.5$ are identical? Explain.

4. (25 pts) **Dummy Variables.** Consider two hypotheses: (a) men make more than women on average and (b) men progress up the pay scale faster than women as they age. Modify equation (1) as necessary and describe the how to test each hypothesis.

5. (25 pts) **Spline Regression**. Derive the spline regression model with a knot at $Age = 40$.

6. (25 pts) **Marginal Effects**. Consider a modified version of equation (1) with a quadratic age-earnings profile. Describe a test procedure to test that $\frac{\partial wage}{\partial age} = 0.5$.

7. (50 pts) **Maximum Likelihood Estimation**. Assume the error terms have beta($\alpha = 2, \gamma = 1$) distribution:

$$f(\epsilon_t; c) = c\epsilon_t^{\alpha-1}(1-\epsilon_t)^{\gamma-1},$$

where $0 \leq \epsilon_t \leq 1$.

(a) Find the constant c.

(b) How does the new error distribution change your answer in part (1)?

(c) Find the first- and second-order conditions for maximizing the log likelihood function.

(d) Calculate the Newton-Raphson and steepest-ascent algorithms for maximizing the log likehood function.

(e) What is the maximum likelihood estimator under normally distributed errors?